

*** The Effects of Tablet-Based Electronic Grading on Airline Evaluator Performance ***

Author: M. Elsenrath

Parks College of Engineering, Aviation and Technology, Saint Louis University, USA

Contact Information:

mike.elsenrath@slu.edu

General Introduction

The contents of this dataset include all quantitative data collected in association with the Ph.D. dissertation research of Michael Elsenrath (<https://search.proquest.com/docview/2308668820>). Most data were collected at a major U.S. airline's training facility, with a small number of trials occurring at offsite locations. Data collection occurred from November of 2018, through April of 2019. The data is being made public as a supplement to published work, and to allow replication and advancement of the author's research.

Purpose

The purpose of the research was to investigate the effect of tablet-based grading formats on airline evaluator performance to improve industry safety and efficiency.

Test Equipment

Two video recordings were made of an airline flight crew experiencing an engine failure during departure in a Boeing 767 full motion flight simulator. Video files included relevant information regarding the flight crew's performance (e.g., airspeed, attitude, etc.). All research participants viewed the videos using a 32" high definition monitor and graded student performance. Various grading formats were assessed to determine which format (i.e., pen and paper, iPad paper equivalent, or the iPad) provided the closest alignment with the referent standard score in terms of both rater agreement and correlation measures.

Description of the data in the dataset

Two Microsoft Excel files are included. Most of the requisite information required to understand the nature of the data is contained within each spreadsheet. The following provides amplifying information and is broken down by file name.

File Name: iPad Psychometric Validity and Reliability Data

This file contains the data collected in association with the psychometric testing of the iPad grading software used in the research. The iPad grading program developed for the research required assurance of its validity and reliability from a psychometric perspective. Seven spreadsheets reside within this file. The first spreadsheet contains ordinal data collected to determine the validity of the grading program. The Content Validity Index was selected from the literature as the validity assessment methodology. The spreadsheet contains relevant equations and references for reproduction of the work.

Spreadsheets two through four contain test-retest correlations of rater assigned grades as a measure of the reliability of the grade sheet. Spreadsheets five through seven contain the criteria

selected during the test and retest that underlie the assigned grade. Each spreadsheet contains pertinent information relative to the statistical tests used for assessment.

File Name: Experimental Data

This file contains the data collected during the main research phase. Four spreadsheets are contained within the file. The first spreadsheet contains the count of criteria selected by each research participant, broken down by experimental group. Differences between groups were assessed using the Kruskal-Wallis test.

Spreadsheets two through four contain the grade assignments made by each rater within each group, as well as the referent true performance score (each spreadsheet represents a grading methodology as highlighted on the Excel sheet tab). Rater performance was assessed between each rater and the referent score using rater agreement and correlation measures to derive an individual rater performance score. Once an individual performance score was established, group comparisons occurred. Refer to the information contained on each sheet regarding the statistics used.